

## Unit 6

**Q. Explain the basic concept of Parallel Processing? 6 M S-17**

**Q. Explain the classification of parallel architecture? 7 M S-17**

**Q. What is the need of parallel processing? Explain the classification of parallel architecture. 7M S-18**

**Ans:**

- Parallel Processing is a processing in which many calculations or execution of processes is carried out simultaneously.
- **Parallel Processing Systems** are designed to speed up the execution of programs by dividing the program into multiple fragments and processing these fragments simultaneously.
- Such systems are multiprocessor systems also known as tightly coupled systems.
- Parallel systems deal with the simultaneous use of multiple computer resources that can include a single computer with multiple processors, a number of computers connected by a network to form a parallel processing cluster or a combination of both.
- Parallel computing is an evolution of serial computing where the jobs are broken into discrete parts that can be executed concurrently.
- Each part is further broken down to a series of instructions. Instructions from each part execute simultaneously on different CPUs.
- Flynn has classified the computer systems based on parallelism in the instructions and in the data streams.

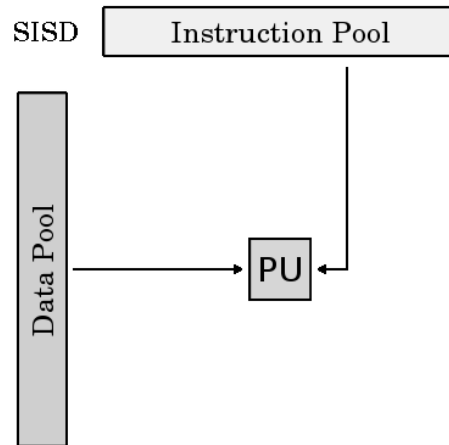
These are:

1. Single instruction stream, single data stream (SISD).
2. Single instruction stream, multiple data stream (SIMD).
3. Multiple instruction streams, single data stream (MISD).
4. Multiple instruction stream, multiple data stream (MIMD).

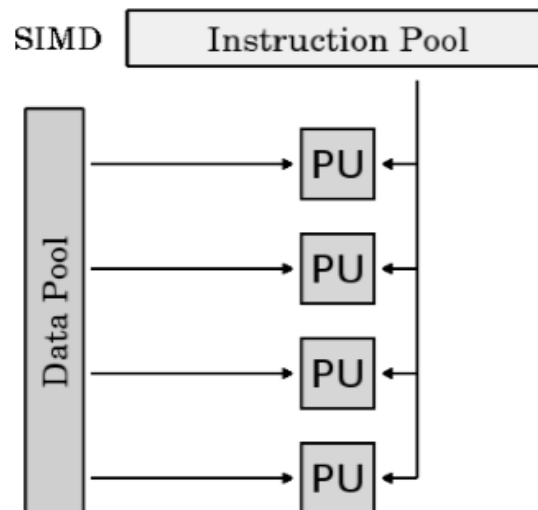
**1 Single instruction stream, single data stream (SISD).**

- A serial (non-parallel) computer
- Single Instruction: Only one instruction stream is being acted on by the CPU during any one clock cycle
- Single Data: Only one data stream is being used as input during any one clock cycle
- Deterministic execution

- This is the oldest type of computer
- Examples: older generation mainframes, minicomputers, workstations and single processor/core PCs



## 2. Single instruction stream, multiple data stream (SIMD).

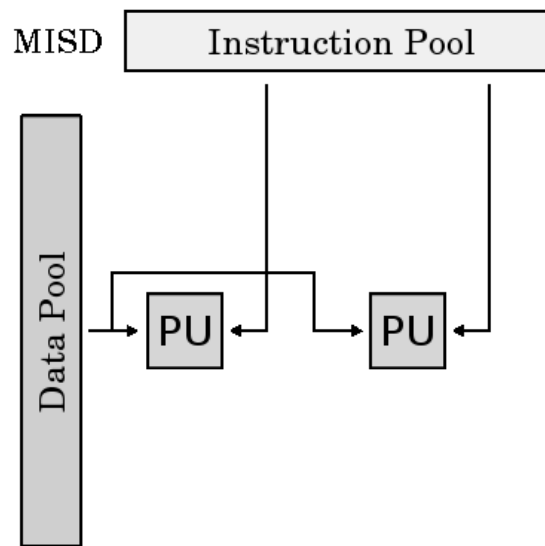


- It represents the organization of a single computer containing a control unit, processor unit and a memory unit. Instructions are executed sequentially
- It can be achieved by pipelining or multiple functional units.
- It represents an organization that includes multiple processing units under the control of a common control unit.
- All processors receive the same instruction from control unit but operate on different parts of the data.
- They are highly specialized computers.
- They are basically used for numerical problems that are expressed in the form of vector or matrix. But they are not suitable for other types of computations

A type of parallel computer

- **Single Instruction:** All processing units execute the same instruction at any given clock cycle
- **Multiple Data:** Each processing unit can operate on a different data element
- Best suited for specialized problems characterized by a high degree of regularity, such as graphics/image processing.
- Synchronous (lockstep) and deterministic execution

### 3. Multiple instruction streams, single data stream (MISD).



- It consists of a single computer containing multiple processors connected with multiple control units and a common memory unit.
- It is capable of processing several instructions over single data stream simultaneously.
- MISD structure is only of theoretical interest since no practical system has been constructed using this organization.
- A type of parallel computer
- Multiple Instruction: Each processing unit operates on the data independently via separate instruction streams.

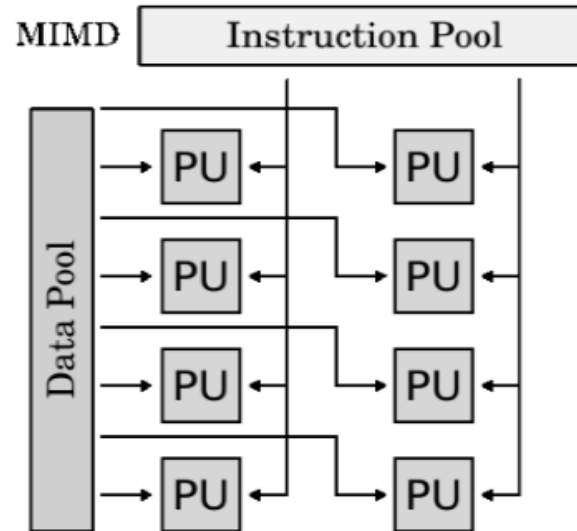
**Single Data:** A single data stream is fed into multiple processing units.

Few (if any) actual examples of this class of parallel computer have ever existed.

Some conceivable uses might be:

- multiple frequency filters operating on a single signal stream
- multiple cryptography algorithms attempting to crack a single coded message.

#### 4. Multiple instruction stream, multiple data stream (MIMD).



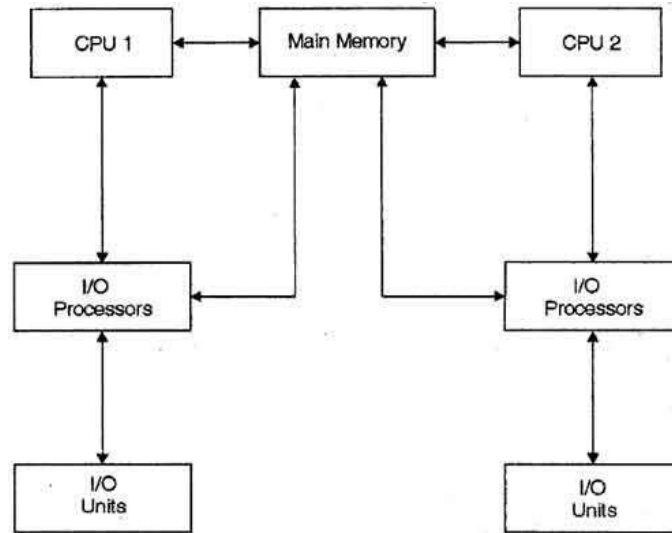
- It represents the organization which is capable of processing several programs at same time.
- It is the organization of a single computer containing multiple processors connected with multiple control units and a shared memory unit. The shared memory unit contains multiple modules to communicate with all processors simultaneously.
- Multiprocessors and multicomputer are the examples of MIMD. It fulfills the demand of large scale computations.

A type of parallel computer

- Multiple Instruction: Every processor may be executing a different instruction stream
- Multiple Data: Every processor may be working with a different data stream
- Execution can be synchronous or asynchronous, deterministic or non-deterministic
- Currently, the most common type of parallel computer - most modern supercomputers fall into this category.
- Examples: most current supercomputers, networked parallel computer clusters and "grids", multi-processor SMP computers, multi-core PCs.
- Note: many MIMD architectures also include SIMD execution sub-component

## Multi processor System

Ans:



- **Multiprocessor Operating System** refers to the use of two or more central processing units (CPU) within a single computer system.
- These multiple CPUs are in a close communication sharing the computer bus, memory and other peripheral devices. These systems are referred as *tightly coupled systems*.
- These types of systems are used when very high speed is required to process a large volume of data. These systems are generally used in environment like satellite control, weather forecasting etc. The basic organization of multiprocessing system is shown in fig.
- Multiprocessing system is based on the symmetric multiprocessing model, in which each processor runs an identical copy of [operating system](#) and these copies communicate with each other.
- In this system processor is assigned a specific task. A master processor controls the system. This scheme defines a master-slave relationship.
- These systems can save money in compare to single processor systems because the processors can share peripherals, power supplies and other devices. The main advantage of multiprocessor system is to get more work done in a shorter period of time. Moreover, multiprocessor systems prove more reliable in the situations of failure of one processor. In this situation, the system with multiprocessor will not halt the system; it will only slow it down.

In order to employ multiprocessing operating system effectively, the computer system must have the followings:

**1. Motherboard Support:** A motherboard capable of handling multiple processors. This means additional sockets or slots for the extra chips and a chipset capable of handling the multiprocessing arrangement.

**2. Processor Support:** processors those are capable of being used in a multiprocessing system. The whole task of multiprocessing is managed by the operating system, which allocates different tasks to be performed by the various processors in the system.

### Advantages of Multiprocessor Systems

There are multiple advantages to multiprocessor systems. Some of these are:

#### More reliable Systems

In a multiprocessor system, even if one processor fails, the system will not halt. This ability to continue working despite hardware failure is known as graceful degradation. For example: If there are 5 processors in a multiprocessor system and one of them fails, then also 4 processors are still working. So the system only becomes slower and does not ground to a halt.

#### Enhanced Throughput

If multiple processors are working in tandem, then the throughput of the system increases i.e. number of processes getting executed per unit of time increase. If there are N processors then the throughput increases by an amount just under N.

#### More Economic Systems

Multiprocessor systems are cheaper than single processor systems in the long run because they share the data storage, peripheral devices, power supplies etc. If there are multiple processes that share data, it is better to schedule them on multiprocessor systems with shared data than have different computer systems with multiple copies of the data.

### Disadvantages of Multiprocessor Systems

There are some disadvantages as well to multiprocessor systems. Some of these are:

#### Increased Expense

Even though multiprocessor systems are cheaper in the long run than using multiple computer systems, still they are quite expensive. It is much cheaper to buy a simple single processor system than a multiprocessor system.

#### Complicated Operating System Required

There are multiple processors in a multiprocessor system that share peripherals, memory etc. So, it is much more complicated to schedule processes and impart resources to processes than in single processor systems. Hence, a more complex and complicated operating system is required in multiprocessor systems.

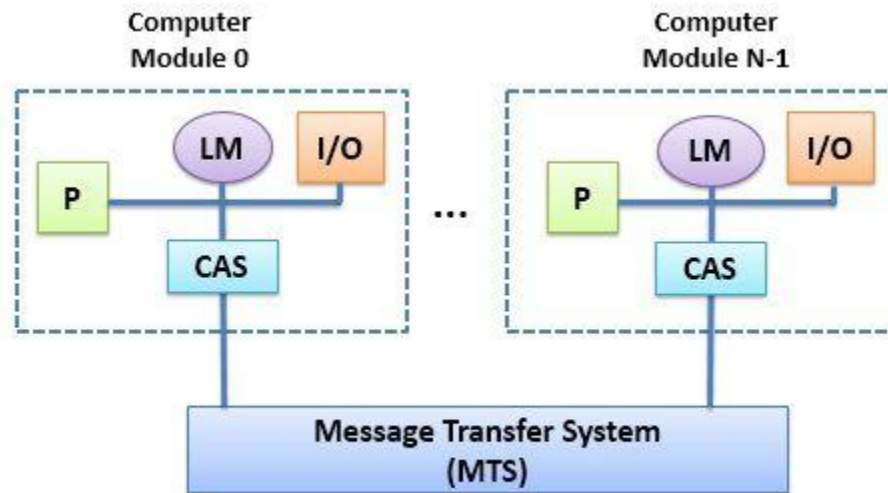
#### Large Main Memory Required

All the processors in the multiprocessor system share the memory. So a much larger pool of memory is required as compared to single processor systems.

**Q. Describe the loosely and tightly coupled multi-computer system. 7 M W-18, W-16**

**Ans: Loosely Coupled Multiprocessor System**

Multiprocessor is one which has more than two processors in the system. Now when the **degree of coupling** between these processors is very **low**, the system is called **loosely coupled multiprocessor system**. In loosely coupled system each processor has its **own local memory, a set of input-output devices** and a **channel and arbiter switch (CAS)**. We refer to the processor with its local memory and set of input-output devices and CAS as a **computer module**.



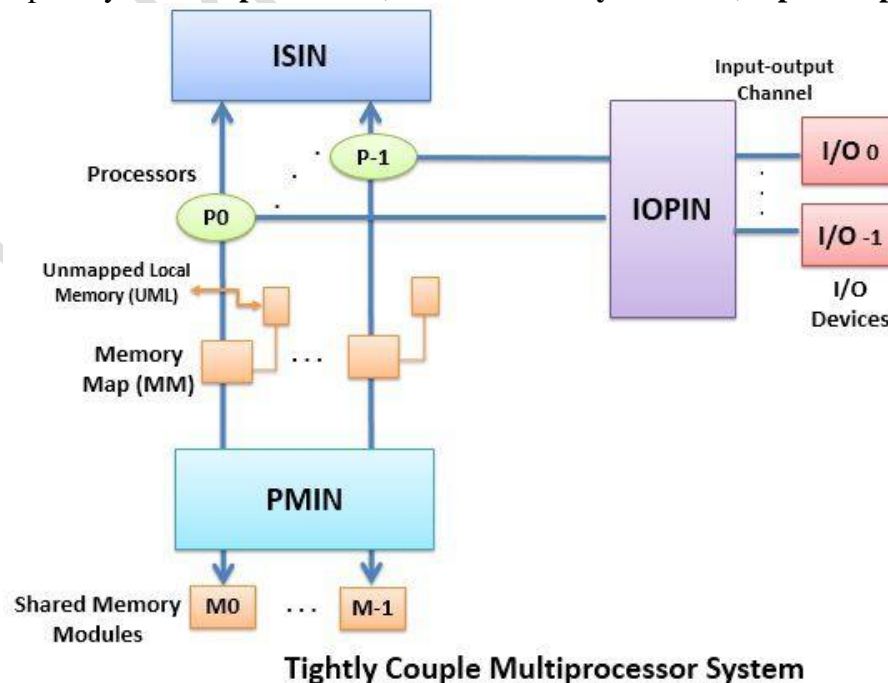
### Loosely Couple Multiprocessor System

Processes that execute on different computer modules communicate with each other by exchanging the **messages** through a physical segment of **message transfer system (MTS)**. The loosely coupled system is also known as **distributed system**. The loosely coupled system is **efficient** when the processes running on different computer module require **minimal interaction**.

If the request for accessing MTS of two or more computer module collide, the **CAS responsibly** chooses one of the simultaneous requests and delay other requests until selected request is serviced completely. The CAS has a **high-speed communication memory** which can be accessed by all the processors in the system. The communication memory in CAS is used to **buffer the transfers of messages**.

### Tightly Coupled Multiprocessor System

The **throughput** of the loosely coupled system may be **too low** for some of the applications that require **fast access time**. In this case, **Tightly coupled microprocessor system** must be used. The tightly coupled system has **processors, shared memory modules, input-output channels**.



### Tightly Couple Multiprocessor System



The above units of the tightly coupled system are connected through the set of three **interconnection networks**, **processor-memory interconnection network (PMIN)**, **I/O-processor interconnection network (IOPIN)** and the **interrupt-signal interconnection network (ISIN)**. The use of these three interconnection networks is as follow.

**PMIN:** It is a switch which **connects** each **processor** to every **memory module**. It can also be designed in a way that a processor can broadcast data to one or more memory module.

**ISIN:** It allows each **processor** to **direct an interrupt** to any **other processor**.

**IOPIN:** It allows a **processor** to **communicate** with an **I/O channel** which is connected to input-output devices.

#### Q. Difference Between Loosely Coupled & Tightly Coupled Multiprocessor System

Comparison	Loosely Coupled Multiprocessor System	Tightly Coupled Multiprocessor System
Basic	Each processor has its own memory module.	Processors have shared memory modules.
Efficient	Efficient when tasks running on different processors, has minimal interaction.	Efficient for high-speed or real-time processing.
Memory conflict	It generally, do not encounter memory conflict.	It experiences more memory conflicts.
Interconnections	Message transfer system (MTS).	Interconnection networks PMIN, IOPIN, ISIN.
Data rate	Low.	High.
Expensive	Less expensive.	More expensive.
Cost	The cost of loosely coupled multiprocessor system is less.	Tightly coupled multiprocessor system is more costly.
	In loosely coupled multiprocessor system, modules are connected through <b>Message transfer system</b> network.	While there is PMIN, IOPIN and ISIN networks.

#### Q. Draw and explain the uniform and non-uniform memory access multiprocessor system.

7 M W-18, W-17

Ans:

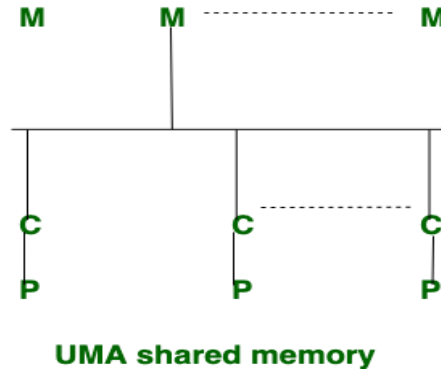
Multiprocessors can be categorized into two shared-memory model which are:

1. Uniform Memory Access (UMA)
2. Non-uniform Memory Access (NUMA)

#### Uniform Memory Access (UMA):

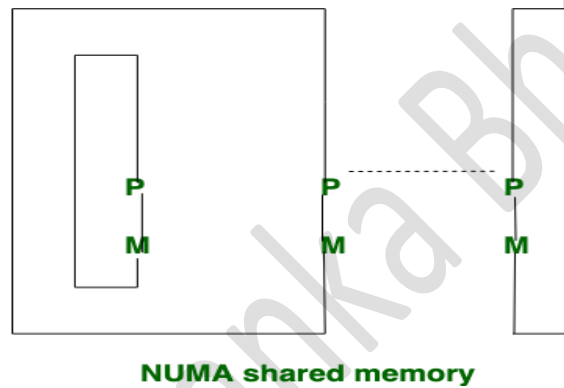
In UMA, where Single memory controller is used. Uniform Memory Access is slower than non-uniform Memory Access. In Uniform Memory Access, bandwidth is restricted or limited rather than non-uniform memory access. There are 3 types of buses used in uniform Memory Access which are: Single, Multiple and Crossbar. It is applicable for general purpose applications and time-sharing applications.





**Non-uniform Memory Access (NUMA):**

In NUMA, where different memory controller is used. Non-uniform Memory Access is faster than uniform Memory Access. Non-uniform Memory Access is applicable for real-time applications and time-critical applications.



Let's see the difference between UMA and NUMA:

S.NO	UMA	NUMA
1.	UMA stands for Uniform Memory Access.	NUMA stands for Non-uniform Memory Access.
2.	In Uniform Memory Access, Single memory controller is used.	In Non-uniform Memory Access, Different memory controller is used.
3.	Uniform Memory Access is slower than non-uniform Memory Access.	Non-uniform Memory Access is faster than uniform Memory Access.
4.	Uniform Memory Access has limited bandwidth.	Non-uniform Memory Access has more bandwidth than uniform Memory Access.
5.	Uniform Memory Access is applicable for general purpose applications and time-sharing applications.	Non-uniform Memory Access is applicable for real-time applications and time-critical applications.
6.	In uniform Memory Access, memory access time is balanced or equal.	In non-uniform Memory Access, memory access time is not equal.
7.	There are 3 types of buses used in uniform Memory Access which are: Single, Multiple and Crossbar.	While in non-uniform Memory Access, There are 2 types of buses used which are: Tree and hierarchical.

**Q. Write a short note on array processor. 7 M W-18, W-16**

**Q. Explain Array processor in detail? 7 M S-17**

**Ans:**

Array processors are also known as multiprocessors. They perform computations on large arrays of data. Thus, they are used to improve the performance of the computer.

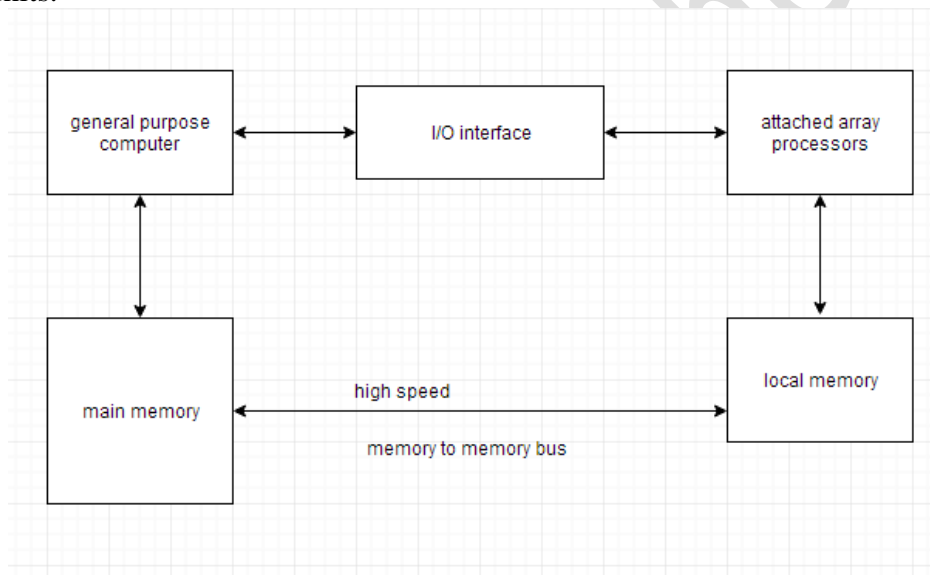
### Types of Array Processors

There are basically two types of array processors:

1. Attached Array Processors
2. SIMD Array Processors

### Attached Array Processors

An attached array processor is a processor which is attached to a general purpose computer and its purpose is to enhance and improve the performance of that computer in numerical computational tasks. It achieves high performance by means of parallel processing with multiple functional units.



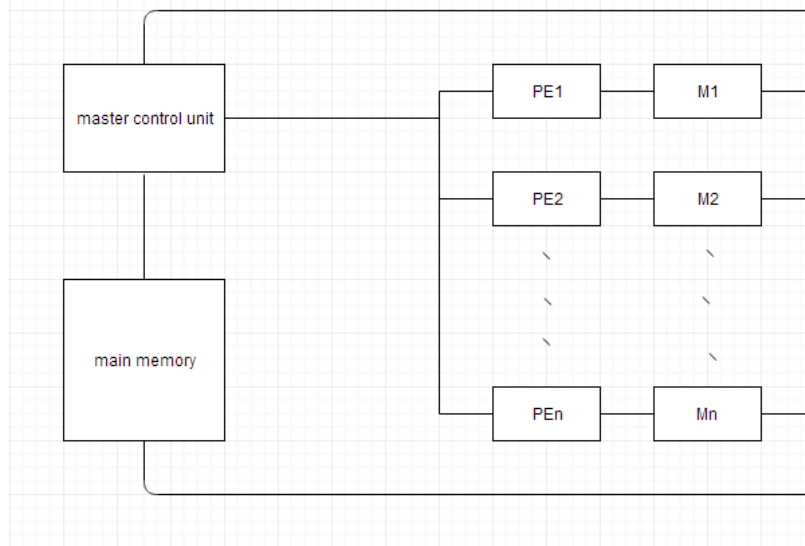
### SIMD Array Processors

SIMD is the organization of a single computer containing multiple processors operating in parallel. The processing units are made to operate under the control of a common control unit, thus providing a single instruction stream and multiple data streams.

A general block diagram of an array processor is shown below. It contains a set of identical processing elements (PE's), each of which is having a local memory M. Each processor element includes an **ALU** and **registers**. The master control unit controls all the operations of the processor elements. It also decodes the instructions and determines how the instruction is to be executed.

The main memory is used for storing the program. The control unit is responsible for fetching the instructions. Vector instructions are sent to all PE's simultaneously and results are returned to the memory.

The best known SIMD array processor is the **ILLIAC IV** computer developed by the **Burroughs corps**. SIMD processors are highly specialized computers. They are only suitable for numerical problems that can be expressed in vector or matrix form and they are not suitable for other types of computations.

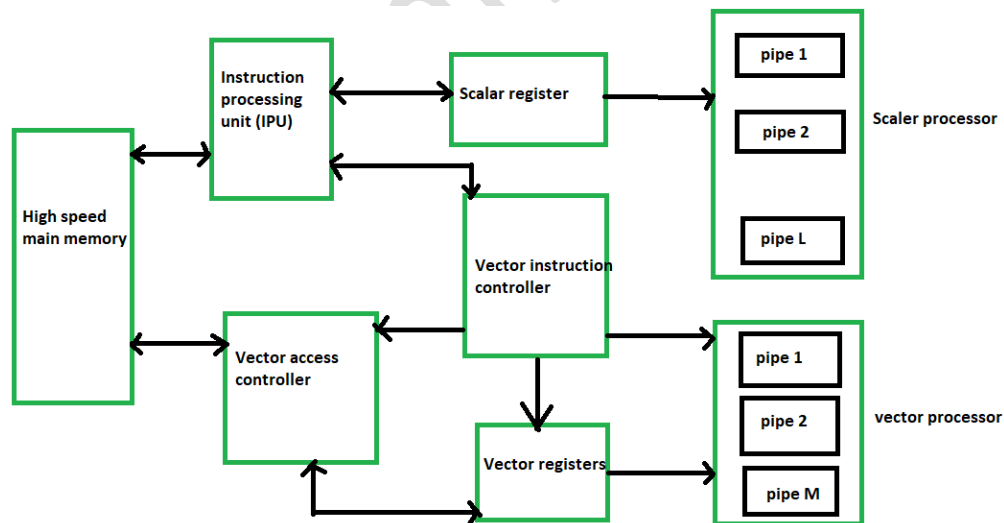


### Why use the Array Processor

- Array processors increase the overall instruction processing speed.
- As most of the Array processors operate asynchronously from the host CPU, hence it improves the overall capacity of the system.
- Array Processors has its own local memory, hence providing extra memory for systems with low memory.

### Q. Explain vector processor with suitable example. 7 M W-18

**Ans:** A block diagram of a modern multiple pipeline vector computer is shown below:



According to from where the operands are retrieved in a vector processor, pipe lined vector computers are classified into two architectural configurations:

1. **Memory to memory architecture** –In memory to memory architecture, source operands, intermediate and final results are retrieved (read) directly from the main memory. For memory to memory vector instructions, the information of the base address, the offset, the increment, and the the vector length must be specified in order to enable streams of data transfers between the main memory and pipelines. The processors like *TI-ASC*, *CDC*

*STAR-100, and Cyber-205* have vector instructions in memory to memory formats. The main points about memory to memory architecture are:

- There is no limitation of size
- Speed is comparatively slow in this architecture

**2. Register to register architecture –**

In register to register architecture, operands and results are retrieved indirectly from the main memory through the use of large number of vector registers or scalar registers. The processors like *VP-200* use vector instructions in register to register formats. The main points about register to register architecture are:

- Register to register architecture has limited size.
- Speed is very high as compared to the memory to memory architecture.
- The hardware cost is high in this architecture.

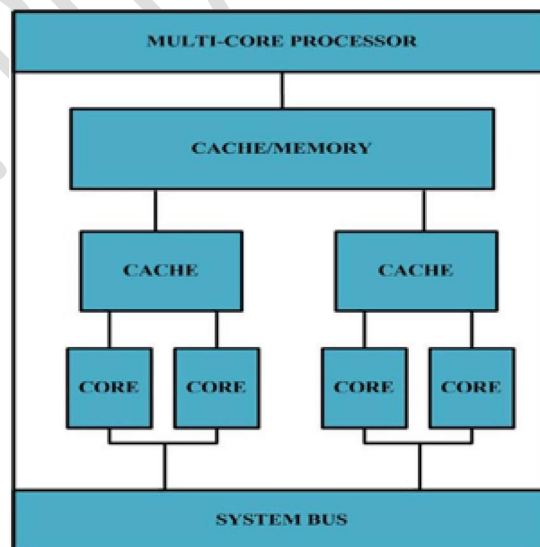
**Q. What is multicore architecture why there is a need of multicore architecture? 6M S-18**

**Q. What is multicore architecture explain in detail? 6 M S-17**

**Ans:** A multi-core architecture (or a chip multiprocessor) is a general-purpose processor that consists of multiple cores on the same die and can execute programs simultaneously.

Multicore refers to an architecture in which a single physical processor incorporates the core logic of more than one processor. A single integrated circuit is used to package or hold these processors. These single integrated circuits are known as a die. Multicore architecture places multiple processor cores and bundles them as a single physical processor. The objective is to create a system that can complete more tasks at the same time.

This technology is most commonly used in multicore processors, where two or more processor chips or cores run concurrently as a single system. Multicore-based processors are used in mobile devices, desktops, workstations.



The concept of multicore technology is mainly centered on the possibility of parallel computing, which can significantly boost computer speed and efficiency by including two or more central processing units (CPUs) in a single chip. This reduces the system's heat and power consumption. This means much better performance with less or the same amount of energy.

The architecture of a multicore processor enables communication between all available cores to ensure that the processing tasks are divided and assigned accurately. At the time of task completion, the processed data from each core is delivered back to the motherboard by means of a single shared gateway. This technique significantly enhances performance compared to a single-core processor of similar speed

**Q. Draw and explain single bus inter connection network. 6 M W-17, W-16**

**Q. Draw and explain cross bar inter connection network. 6 M W-16**

**Ans:** Interconnection network are used to connect nodes, where nodes can be a single processor or group of processors, to other nodes.

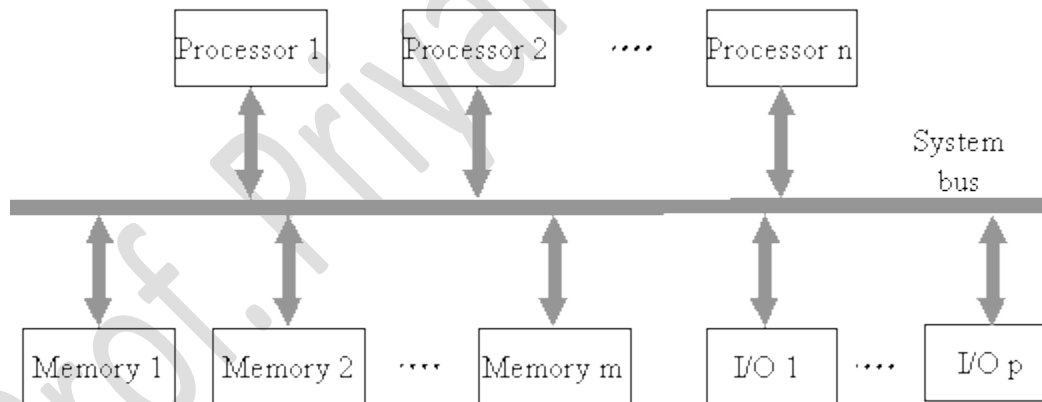
Interconnection networks can be categorized on the basis of their topology. Topology is the pattern in which one node is connected to other nodes.

Static interconnection networks for elements of parallel systems (ex. processors, memories) are based on fixed connections that can not be modified without a physical re-designing of a system. Static interconnection networks can have many structures such as a linear structure (pipeline), a matrix, a ring, a torus, a **complete connection** structure, a tree, a star, a **hyper-cube**.

Dynamic interconnection networks between processors enable changing (reconfiguring) of the connection structure in a system. It can be done before or during parallel program execution. So, we can speak about **static** or **dynamic connection reconfiguration**.

### Single Bus Interconnection Network

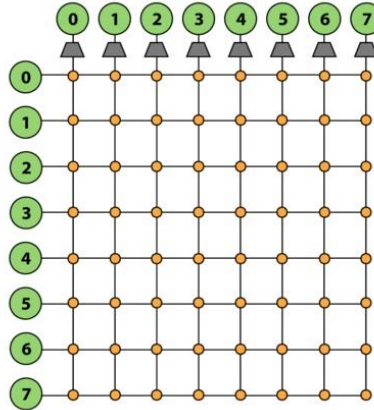
A bus is the simplest type of dynamic interconnection networks. It constitutes a common data transfer path for many devices. Depending on the type of implemented transmissions we have **serial busses** and **parallel busses**. The devices connected to a bus can be processors, memories, I/O units, as shown in the figure below.



Only one device connected to a bus can transmit data. Many devices can receive data. In the last case we speak about a **multicast transmission**. If data are meant for all devices connected to a bus we speak about a **broadcast transmission**. Accessing the bus must be synchronized. It is done with the use of two methods: a **token method** and a **bus arbiter method**. With the token method, a token (a special control message or signal) is circulating between the devices connected to a bus and it gives the right to transmit to the bus to a single device at a time. The bus arbiter receives data transmission requests from the devices connected to a bus. It selects one device according to a selected strategy (ex. using a system of assigned priorities) and sends an acknowledge message (signal) to one of the requesting devices that grants it the transmitting right. After the selected device completes the transmission, it informs the arbiter that can select another request. The receiver (s) address is usually given in the header of the message. Special header values are used for the broadcast and multicasts. All receivers read and decode headers. These devices that are specified in the header, read-in the data transmitted over the bus.

### Cross Bar Inter Connection Network

A crossbar switch is a circuit that enables many interconnections between elements of a parallel system at a time. A crossbar switch has a number of input and output data pins and a number of control pins. In response to control instructions set to its control input, the crossbar switch implements a stable connection of a determined input with a determined output. The diagrams of a typical crossbar switch are shown in the figure below.



Control instructions can request reading the state of specified input and output pins i.e. their current connections in a crossbar switch. Crossbar switches are built with the use of multiplexer circuits, controlled by latch registers, which are set by control instructions. Crossbar switches implement direct, single **non-blocking connections**, but on the condition that the necessary input and output pins of the switch are free. The connections between free pins can always be implemented independently on the status of other connections. New connections can be set during data transmissions through other connections. The non-blocking connections are a big advantage of crossbar switches. Some crossbar switches enable broadcast transmissions but in a blocking manner for all other connections. The disadvantage of crossbar switches is that extending their size, in the sense of the number of input/output pins, is costly in terms of hardware. Because of that, crossbar switches are built up to the size of 100 input/output pins.